Accurate Rough Terrain Modeling from Fused 3D Point Cloud Data

Mahesh Kr Singh*, K. S. Venkatesh† and Ashish Dutta‡
Department of Electrical Engineering *†, Department of Mechanical Engineering ‡
Indian Institute of Technology Kanpur, India
Email: {ksmahesh, venkats, adutta}@iitk.ac.in

***Abstract*—In this paper, we present a new technique for 3D data fusion from two heterogeneous range acquisition devices (i.e., Laser range scanner and Microsoft Kinect) for the extraction of accurate, realistic and rapidly surface reconstruction. First, we present an unsupervised classification algorithm to classify the 3D point cloud data of the terrain into coarser and finer regions. The classification of the 3D point cloud data is done by exploiting the statistical measurement properties of the range dataset. The main merits of the classification method are threshold-freedom and independence from 3D data format and resolution, while preserving characteristic of the terrain details. The 3D point cloud acquired from both the range scanners is transformed into a common reference frame using the principle component algorithm. In the reference frame, the fused point cloud data are obtained by integration of coarser regions data from Kinect and finer regions data from a Laser range scanner. The fused point cloud data eliminate the demerits of the both range sensors by complementing each other. After fusion, we apply Delaunay triangulation algorithm to generate the highly accurate, realistic 3D surface of the terrain. Finally, the experimental results demonstrate the highly robust and precision of the proposed approach.**

Keywords - Laser range scanner; Unsupervised segmentation; Point cloud model; Principle component analysis; Delaunay triangulation.

I. INTRODUCTION

The multi-range sensor data fusion is the process of combining the 3D information from, redundant and/or complementary sensors, to produce a complete and accurate representation of the targeting environment. The 3D point cloud data fusion has the special significance for the generation of 3D model of the terrain where a large amounts of 3D data must be incorporated and distilled to obtain the best quality terrain information. Nowadays, the generation of dense 3D data to represent the environment has gained more attention. The different kinds of range acquisition system are used to acquire the rough 3D point cloud data to reconstruct the 3D model of the environment such as: Laser Range Scanner (LRS) over pan and/or tilt platforms, stereo-vision systems, Time-of-Flight (ToF) camera system, and more recently the use of RGB-D cameras, like the Microsoft Kinect sensor. These range sensors are competent to measure the detailed depth information of the environment efficiently, but each kind of range sensor has its own advantages and disadvantages. These limitations make them more suitable for the specific kinds of application. So the accurate and efficient 3D modeling of terrain is a challenging task. To build a dense geometrical 3D model of the terrain, different range sensors could be chosen in order to acquire 3D point cloud data. Therefore, fusion of sensory information is essential for the generation of a dense 3D model of the terrain.

In recent years, the generating a seamless integration of surface from multiple overlapping 3D point cloud data have been studied extensively. Some of the first work emphasis on the fusion of point cloud data by making an implicit function [1] and then polygonizing it using the marching cubes algorithm for high resolution surface reconstruction [2]. These algorithms were implemented using a same kind data structure. Hilton and Illingworth [3] have presented a multiresolution surface fusion algorithm. The proposed algorithm combines and compresses data using an octree, but it does not model the sensor noise explicitly. Also, these algorithms do not produce the adaptive resolution surfaces, although the method has straightforward using their corresponding marching triangle algorithm [4]. Wurm et al. [5] have done a proper review of the previous methods, and proposed a technique for modeling 3D plane based on octrees using the probabilistic occupancy estimation. Their approach is able to represent full 3D models of plane including free and unknown sites. Trevor et al. [6] have proposed the SLAM techniques which uses planar surfaces as landmarks, and maps its positions and extent for 3D modeling. In their work, They have used a 3D range scanner (a tilting LRF or a Kinect type sensor) to obtain 3D planar surfaces, that combined with 2D segments acquired from the 2D scanner at the base of a mobile robot, have used to construct a map using the GTSAM library [7]. Therefore, the combination of 2D lines and 3D planes with a high level representation and easy to be annotated with semantic data have generated a precise map with its high level features. An Su et.al. [8] have proposed the fast planar surface detection using 2D lines extracted from a tilting the LRF over mobile robot. The proposed method works in (online) real-time, and only stores initial and end point of each 2D line to create the 3D model. Klaess et al. [9] have built the 3D surface element grid maps and proposed Monte Carlo localization with the probabilistic observation models for 2D and 3D sensors on this map. Rusu et al. [10] have presented the pan rotating laser range finder (LRF) that has used to get a point cloud, also it has used to obtain a high level semantic model like a kitchen environment. The model is generated off-line and used a machine learning technique to codify objects

and labeling them with its semantic information. A tilting LRF [11] has been used to obtain the 3D point cloud data and the machine learning techniques have been applied to distinguish the environment to navigable and non navigable zones. Douillard et al. [12] have used a LRF to build a hybrid 3D outdoor model of the environment using planar faces and elevation level. Singh et al. [13] have presented a new method for range data fusion from two heterogeneous range scanners. They have exploited the terrain characteristic (i.e. coarser and finer region) for fusion of range data and generated accurate 3D fused surface of the planner environment. So many works have used range sensors to model the 3D objects and for surface generation [14], [15], the modeled objects have used to construct semantic maps or for the pattern and object recognition, have applications for the color image and depth recognition or to help the robot to recognize and grasp different objects.

In this paper, we present an unsupervised classification algorithm to classify the coarser and finer regions of the environment. The classification of the 3D point cloud data is done by exploiting the statistical measurement properties of the point cloud dataset. After classification, we determine the location of finer regions of the terrain. The 3D point clouds acquired from both the range scanners are transformed into a common reference frame using the principle component analysis method. In the reference frame, the fused point cloud data are obtained by the integration of coarser regions data from Kinect and finer regions data from a Laser range scanner. The fused point cloud data eliminate the demerits of the both range sensors by complementing each other. After fusion, we apply Delaunay triangulation algorithm to generate the highly accurate, realistic 3D surface of the terrain.

The remainder of this paper is organized as follows: Section II describes the proposed method in detail. Section III presents the experimental results. Finally, we conclude this paper in Section IV.

II. Proposed Method

In this section, we briefly describe a new fusion method of two heterogeneous range acquisition systems for Accurate, realistic, and fast 3D representation of the terrain. The steps are detailed in the following:

A. Range data acquisition systems

For the fusion of range data, we have used two heterogeneous range sensors i.e. Laser range scanner and Microsoft Kinect [16], [17]. The Figure 1(a) shows the Laser range scanner which is designed at our Robotics Lab. The 3D Laser scanner is made by an electromechanical devices: a CCD camera, a Diode Red Laser as a line projector, cylindrical lens, a bipolar stepper motor, a Atmega 16 micro controller, microstepping motor driver (A4988), two XigBee for RX/TX communication. The schematic circuit diagram of range scanner is shown in Fig. 2. In the circuit diagram, the stepper motor is connected to the A4988 microstepper motor. The A4988 microstepper motor driver is connected to the $\mathbf{B}^{th}$ port of the ATMega-16L microcontroller. The wireless connection is established between microcontroller and computer through XigBee. When we send the rotation command to range scanner from the computer, the stepper motor rotates accordingly. The Laser projects laser dot which is converted to the laser line through a cylindrical lens on the terrain and camera captures the laser line profile. When range scanner moves over the object surface, the camera acquires images of the distorted pattern which is reflected by the object surface with respect to reference pattern. The height of the object is obtained by taking into account the distortion of the laser light stripe caused by its shape. The designed Laser range sensor gives the accurate range measurements of large angular field of the terrain with angular resolution 0.1125^0. The Laser range scanner produces dense high-resolution 3D point cloud data of the environment. At the end, we assemble different sampled line profile in the common coordinate system to generate a 3D map of the scanned surface. The cost of the Laser range scanner is approximately 400 USD. The accuracy of the range scanner is approximately $\pm$2-4 mm throughout its range. The total range cover of the designed scanner is 250 cm but it can be increased as increase the viewing region of the camera. The major advantages of Laser range scanner: it gives accurate result, very high angular resolution, no correspondence issue because the camera acquires the illuminated scene to obtain the dense 3D geometric information in a single exposure. The disadvantages of the scanner is its slow scanning time due to its hardware constraints (i.e. fixed set-up). The Fig. 1(b) shows the Kinect sensor that was introduced in Nov 2010 by Microsoft for the Xbox-360 video game system. It consists of an infrared projector, an infrared camera and an RGB camera. Microsoft Kinect depth measurement is based on a triangulation methodology. The detail description of Microsoft Kinect has described in the article [16], [18]. Besides using it to map the 3D environment, Kinect is used alongside with the inertial measurement unit to give the position and orientation information to the scanner. Khoshelhem et al. [16] have investigated the accuracy and resolution of Kinect depth data for indoor mapping applications. The Authors have demonstrated that the random error of depth measurement increase quadratically with increasing the distance from the sensor and it ranges from a few millimeters up to 4 cm at the maximum range of the sensor. The depth resolution is also decreased quadratically with increasing distance from the Kinect sensor. At the maximum range, the point spacing in depth along the optical axis of the Kinect sensor is more than 7 cm. For the mapping application, the working range should be within 1-3 meter distance from the sensor otherwise the quality of data is deteriorated by noise and low resolution.

B. Segmentation Methodology

This section deals with separation of coarser region points from finer region points of the 3D point cloud data. The central limit theorem [19] states that naturally measured samples will lead to a normal distribution. The assumption is also made that the finer region points may disturb the normal distribution,

(a) (b)

Fig. 1. (a) The Laser range scanner (b) Microsoft Kinect system

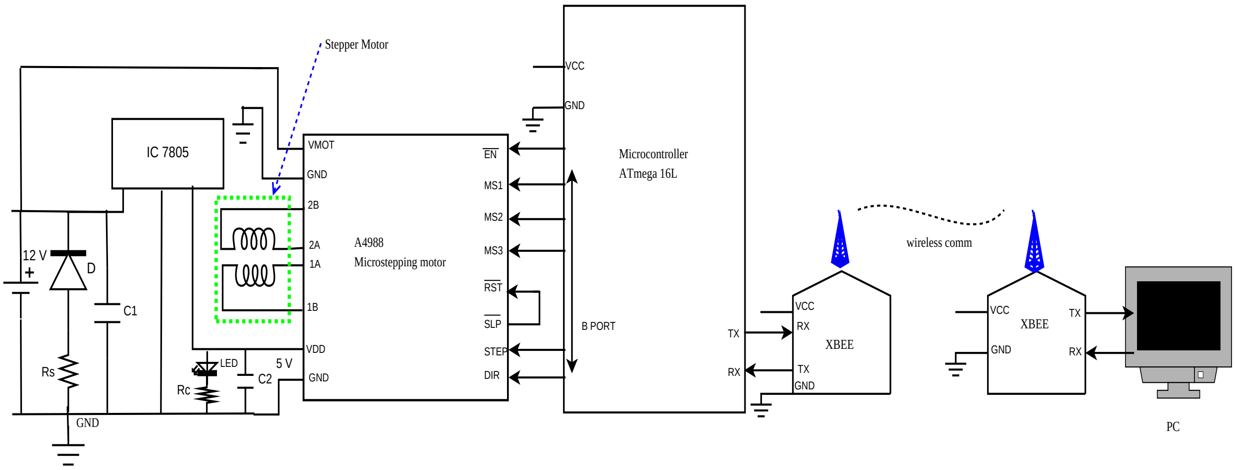


Fig. 2. Schematic circuit diagram of connecting a micro-controller to an A4988 micro-stepping motor driver carrier via XBee to PC.

and by eliminating those points from the 3D point cloud, the coarser region points are obtained. This meaningful statistical measures are needed to describe the 3D point cloud data distribution properly. For a given distribution, skewness S_k [20] is an important asymmetry measure which is a third order moment about mean. Other important measure of distribution is the kurtosis K_u which is a fourth order moment about mean. If the a distribution is symmetric, kurtosis measure the central peak of distribution. The statistic measures are defined as follows:

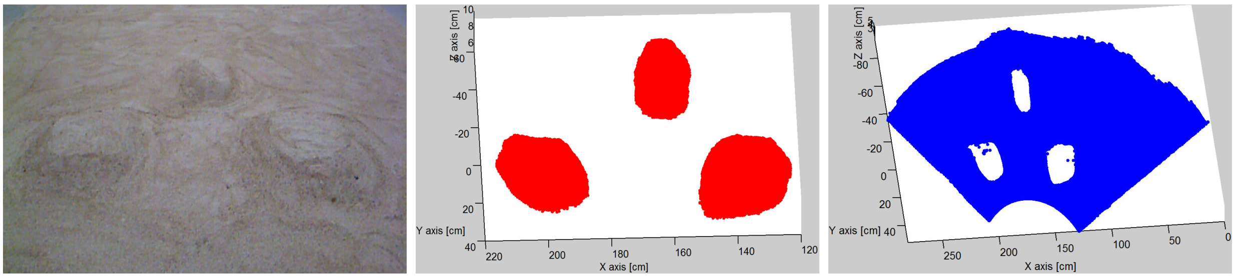
$$\begin{aligned} S_k &= \frac{1}{N\sigma^3}\sum_{i=1}^{N}(r_i-\mu_i)^3 \\ K_u &= \frac{1}{N\sigma^4}\sum_{i=1}^{N}(r_i-\mu_i)^4 \end{aligned} \tag{1}$$

where N is the total number of the range data points r_i with $i \in 1,2,3,4....N$, σ the standard deviation and μ_i the arithmetic mean. As shown in Table 1, for a normal distribution, S_k is zero and K_u is three; if peaks dominate in a point cloud, S_k is greater than zero and K_u is greater than three; if a point cloud is characterised by valleys, S_k is less than zero and K_u is less than three. The segmentation method works as follows:

TABLE I
MEASURES OF DISTRIBUTION

Characteristic of distribution	Normal distribution	Dominance of peaks	Dominance of valleys
Skewness	$S_k = 0$	$S_k > 0$	$S_k < 0$
Kurtosis	$K_u = 3$	$K_u > 3$	$K_u < 3$

The skewness and the kurtosis both demonstrate the characteristics of the 3D point cloud distribution , they can equally be treated as termination criteria in a segmentation algorithm. In the proposed unsupervised segmentation method, skewness is taken as a measure to characterize the point cloud distribution. First, the skewness of the range data is calculated. If skewness is greater than zero, peaks dominate in the range data distribution. In this case, the highest value of the range data is removed by classifying it as a finer region points. To separate all finer region points from ground, these steps in iteratively performed till the skewness $\approx$ 0. The remaining points of the range data belong to the coarser region. In this



(a) (b) (c)

Fig. 3. (a) Figure of sandy terrain, (b) The segmented finer region points (c) The segmented coarser region points of the point cloud.

way, we classify the coarser region points from finer region points of the 3D point cloud data.

C. Data Fusion

Since the range data obtain from both the range acquisition systems are in different coordinate system. Therefore, it is necessary to transform both the 3D point cloud data into a single unified data set for the multi-sensor data fusion, called the transformed reference frame. The fused 3D point cloud data acquired from different scanners eliminate the disadvantages of using scanner alone by complementing each other. The Laser range scanner designed at our Robotics Lab, which scans the environment sequentially. Therefore, the time complexity of 3D data acquisition is large . On the other hand, the 3D point cloud data obtain from the Laser range scanner is very dense, high resolution, accurate, i.e. it has 2-4 mm precision throughout the range from the scanner. In the Kinect, the random depth error increases with increasing distance from the sensor that varies from a few millimeters up to 40 mm at the maximum range of the sensor, but Kinect is a low cost compact range sensor and very fast relative to designed Laser range sensor. Therefore, we use Kinect to scan the coarser area of the terrain and for finer detail area is acquired from designed Laser range scanner. In the process of fusion, first we define the new common coordinate system for both the range sensors. Both the coordinate systems are transformed to new reference coordinate system such that the largest variance of the data is defined as first coordinate (i.e. first principle component) and so on i.e. we transform the basis of 3D point cloud data [21]. The Principal component analysis [21] is a statistical procedure concerned to describe the variance structure of a 3D data set, i.e. it allows to find out the principal directions in which the data are varied. Using segmentation mythology on 3D point cloud data, we determine the location of finer regions of terrain in the reference frame. The main advantage of the proposed method is that it does not require a pre-defined threshold to classify the data. Also, the classification algorithm does not incorporate any prior knowledge about the terrain and is independent of the resolution of the 3D point cloud data. We apply the ICP algorithm [22] to align both the data set in a reference frame. Based on the location of finer region, we enable our Laser scanner to scan only finer region of terrain and the coarser regions of terrain data is taken from Kinect sensor. We have integrated both these data set to generate the fused 3D point cloud data set of the terrain. To build the 3D surface, we apply the Delaunay algorithm to the fused 3D point cloud data. In this way, we reconstruct the accurate, realistic, and fast 3D terrain surface.

III. Experimental Results

The proposed fusion method is performed on range data of the real world by creating different types of environment in Robotics Lab. In the experiment, we have used the two heterogeneous Laser range scanner and Kinect as shown in Fig 1. The aim of the 3D data fusion is to generate the accurate, realistic, and a fast 3D surface of the terrain and eliminates the demerits of both range scanners by complementing each other as mentioned in the previous section. The proposed method is coded in C++/ MATLAB 2014a, and all experiments are performed on a computer with Intel(R) Core(TM) i7-2600 CPU, 8GB RAM and windows 7 operating system. In the experiments, two sets of 3D point cloud data of the same environment are recorded using heterogeneous range sensors. Figure 4(a) shows the rough sandy terrain created in the Robotics Lab. Figure 4(b) shows the surface of sandy terrain acquired from Kinect sensor in Kinect frame. The 3D point cloud data acquired from both the range sensors are in different coordinate system. Therefore, both range data are transformed to a common reference frame. Using segmentation methodology, we determine the finer regions of terrain and enable the Laser range scanner to acquire the range data of finer region only. After acquisition of the 3D data, Figure 4(c-d) shows the surface of sandy terrain in a common reference frame. In the reference frame, there is some disparity exist between these two 3D point cloud data. Therefore, we apply the ICP algorithm to align these two 3D point cloud data. In the fusion process, we retain the coarser detailed regions and erase the fine detailed region of Kinect 3D data and combines finer detailed region of 3D data obtained from the Laser range scanner. To reconstruct the surface, the fused 3D point cloud data are converted via a Delaunay filter to generate an irregular triangulated mesh. Figure 4(e) shows the finally fused 3D surface of the terrain. In this experiment, the sensor disparity of Laser range scanner relative to the Kinect in a reference frame is as the rotation matrix **R**=[0.9704 0.2418 0;-0.2418 0.9704 0; 0 0 1.0000]; and translation vector **t**=[119.4117

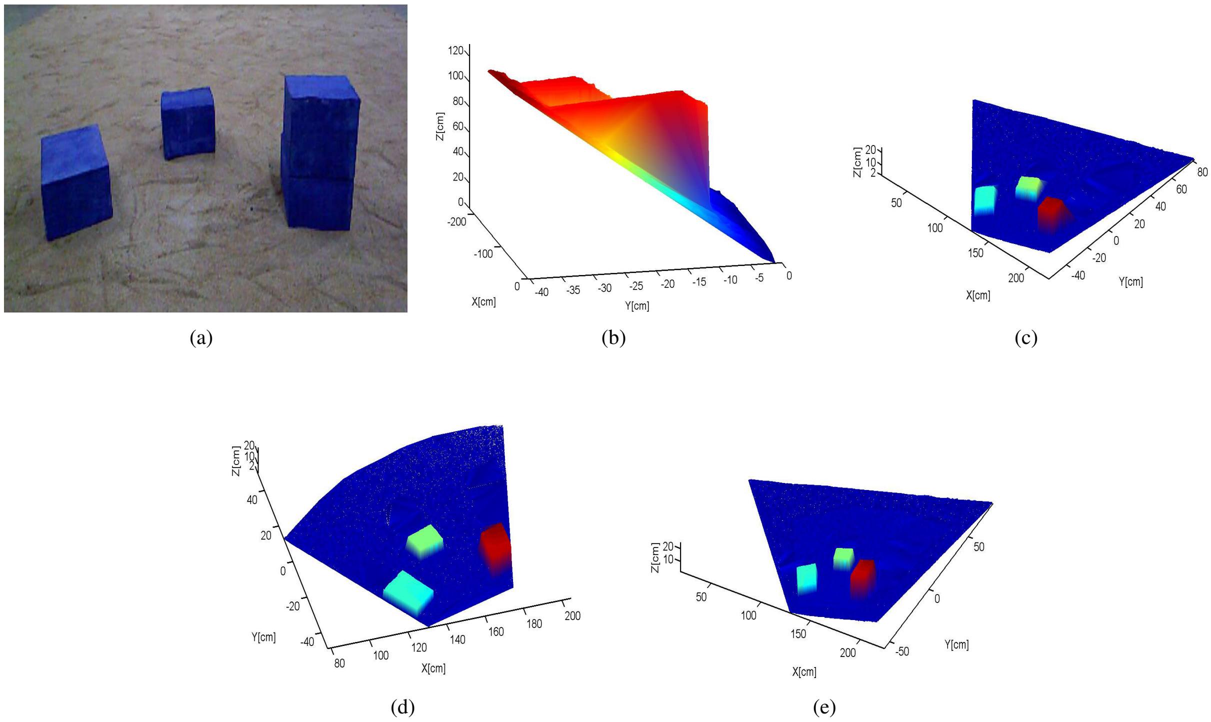


Fig. 4. (a) Figure of different objects are placed on sandy terrain, (b) The surface of the plane in the Kinect frame (c) The surface of the plane in the reference frame (d) The segmented finer region of surface from Laser range scanner (e) The fused 3D surface model from both range sensors.

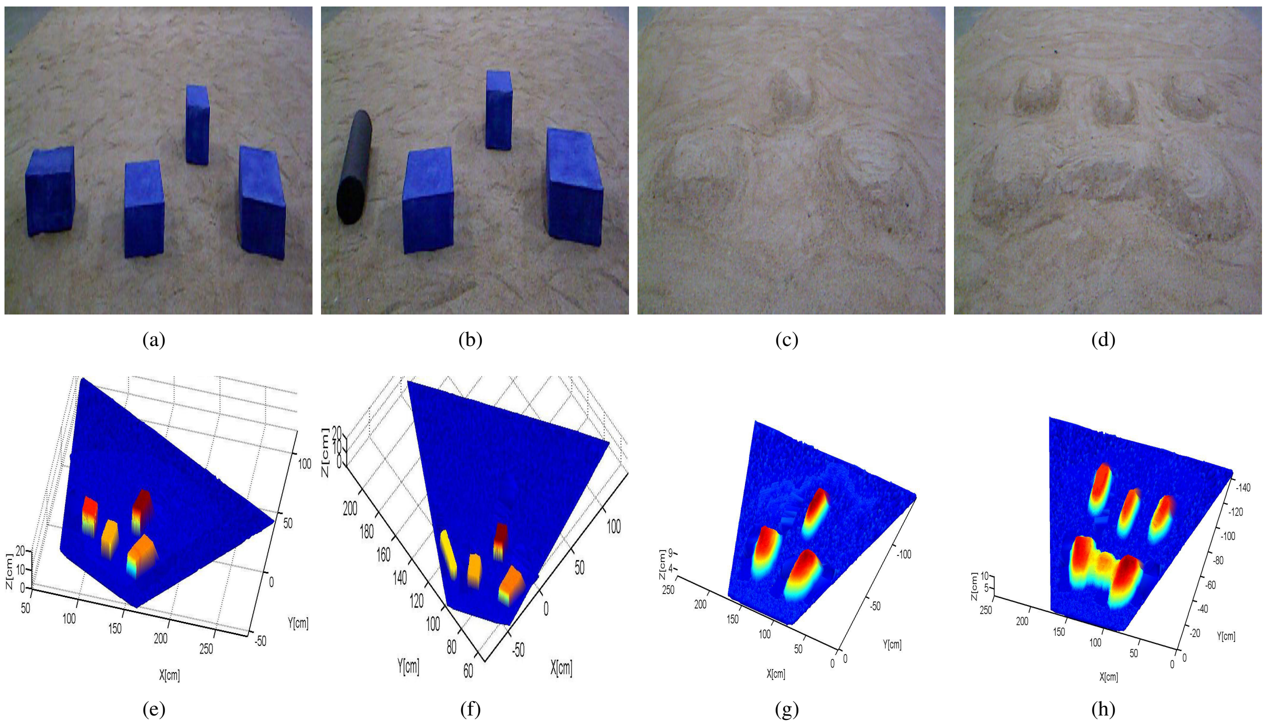


Fig. 5. (a-d) Figure of different kinds of sandy terrain (e-h) The fused accurate 3D model of different kinds of sandy terrain

127.9851 0]. The alignment root mean square error of the fused data is approximately 9.68 mm. Several experiments of point cloud data fusion have been performed in our Robotics laboratory. Fig 5(a-b) shows the different kinds of sandy terrain and their accurate fused 3D terrain model is shown in Fig 5(e-f). Similarly, we created highly rough terrain as shown in Figure 5(c-d). We perform this experiment to check the robustness of the proposed method, i.e. the segmentation of the highly varying finer regions. Fig. 5 (g-h) shows the accurately fused 3D model of the terrain. The resulting 3D fused terrain model shows, the proposed method is efficient to accurately, realistically and rapidly represent the real-world environment.

IV. Conclusions

In this paper, we have presented a new technique for 3D data fusion from two heterogeneous range acquisition devices (i.e. Laser range scanner and Microsoft Kinect) for the extraction of accurate, realistic and rapidly surface reconstruction. First, we have presented an unsupervised classification algorithm for classifying the 3D point cloud data of the terrain into coarser and finer regions. The classification of the 3D point cloud data has been done by exploiting the statistical measurement properties of the dataset. The main merits of the proposed method have a threshold-freedom and independence from 3D data format and resolution, while preserving characteristic of the terrain details. In the reference frame, the fused point cloud data have been obtained by integration of coarser regions data from Kinect and finer regions data from a Laser range scanner. The fused point cloud data have eliminated the demerits of the both range sensors by complementing each other. The 3D fused surface has generated using a Delaunay triangulation algorithm. The experimental results have shown the robustness of proposed method which validated the correctness of the real terrain model. The accurate 3D modeling is used for many applications such as cultural heritage documentation, environment mapping, robot motion and localization, automatic inspection, reverse engineering etc.

References

[1] M. D. Wheeler, Y. Sato, and K. Ikeuchi, "Consensus surfaces for modeling 3d objects from multiple range images," in *Computer Vision, 1998. Sixth International Conference on*. IEEE, 1998, pp. 917–924.

[2] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3d surface construction algorithm," in *ACM Siggraph Computer Graphics*, vol. 21, no. 4. ACM, 1987, pp. 163–169.

[3] A. Hilton and J. Illingworth, "Multi-resolution geometric fusion," in *3-D Digital Imaging and Modeling, 1997. Proceedings., International Conference on Recent Advances in*. IEEE, 1997, pp. 181–188.

[4] A. Hilton, A. J. Stoddart, J. Illingworth, and T. Windeatt, "Marching triangles: range image fusion for complex object modelling," in *Image Processing, 1996. Proceedings., International Conference on*, vol. 1. IEEE, 1996, pp. 381–384.

[5] K. M. Wurm, A. Hornung, M. Bennewitz, C. Stachniss, and W. Burgard, "Octomap: A probabilistic, flexible, and compact 3d map representation for robotic systems," in *Proc. of the ICRA 2010 workshop on best practice in 3D perception and modeling for mobile manipulation*, vol. 2, 2010.

[6] A. J. Trevor, J. Rogers, and H. I. Christensen, "Planar surface slam with 3d and 2d sensors," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*. IEEE, 2012, pp. 3041–3048.

[7] F. Dellaert and M. Kaess, "Square root sam: Simultaneous localization and mapping via square root information smoothing," *The International Journal of Robotics Research*, vol. 25, no. 12, pp. 1181–1203, 2006.

[8] S.-Y. An, L.-K. Lee, and S.-Y. Oh, "Fast incremental 3d plane extraction from a collection of 2d line segments for 3d mapping," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*. IEEE, 2012, pp. 4530–4537.

[9] J. Kläß, J. Stückler, and S. Behnke, "Efficient mobile robot navigation using 3d surfel grid maps," in *Robotics; Proceedings of ROBOTIK 2012; 7th German Conference on*. VDE, 2012, pp. 1–4.

[10] R. B. Rusu, Z. C. Marton, N. Blodow, A. Holzbach, and M. Beetz, "Model-based and learned semantic object labeling in 3d point cloud maps of kitchen environments," in *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*. IEEE, 2009, pp. 3601–3608.

[11] D. F. Wolf and G. Sukhatme, "Semantic mapping using mobile robots," *Robotics, IEEE Transactions on*, vol. 24, no. 2, pp. 245–258, 2008.

[12] B. Douillard, J. Underwood, N. Melkumyan, S. Singh, S. Vasudevan, C. Brunner, and A. Quadros, "Hybrid elevation maps: 3d surface models for segmentation," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. IEEE, 2010, pp. 1532–1538.

[13] M. K. Singh, K. Venkatesh, and A. Dutta, "Accurate 3d terrain modeling by range data fusion from two heterogeneous range scanners," in *India Conference (INDICON), 2014 Annual IEEE*. IEEE, 2014, pp. 1–6.

[14] R. A. Newcombe, A. J. Davison, S. Izadi, P. Kohli, O. Hilliges, J. Shotton, D. Molyneaux, S. Hodges, D. Kim, and A. Fitzgibbon, "Kinectfusion: Real-time dense surface mapping and tracking," in *Mixed and augmented reality (ISMAR), 2011 10th IEEE international symposium on*. IEEE, 2011, pp. 127–136.

[15] K. Lai, L. Bo, X. Ren, and D. Fox, "A large-scale hierarchical multi-view rgb-d object dataset," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1817–1824.

[16] K. Khoshelham and S. O. Elberink, "Accuracy and resolution of kinect depth data for indoor mapping applications," *Sensors*, vol. 12, no. 2, pp. 1437–1454, 2012.

[17] C. Dal Mutto, P. Zanuttigh, and G. M. Cortelazzo, *Time-of-flight cameras and microsoft Kinect*. Springer, 2012.

[18] Z. Zhang, "Microsoft kinect sensor and its effect," *MultiMedia, IEEE*, vol. 19, no. 2, pp. 4–10, 2012.

[19] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification," *Information Theory, IEEE Transactions on*, vol. 13, no. 1, pp. 21–27, 1967.

[20] O. L. Davies and P. L. Goldsmith, *Statistical methods in research and production*. Published for Imperial Chemical Industries by Hafner Pub. Co., 1972.

[21] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2005.

[22] J. Elseberg, S. Magnenat, R. Siegwart, and A. Nüchter, "Comparison of nearest-neighbor-search strategies and implementations for efficient shape registration," *Journal of Software Engineering for Robotics*, vol. 3, no. 1, pp. 2–12, 2012.